Saudi Pharmaceutical Journal xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Saudi Pharmaceutical Journal

journal homepage: www.sciencedirect.com

Original article

Prediction of *Chlamydia pneumoniae* protein localization in host mitochondria and cytoplasm and possible involvements in lung cancer etiology: a computational approach

Aws Alshamsan [a,b,*,1], Shahanavaj Khan [a,*,1], Ahamad Imran [b], Ibrahim A. Aljuffali [a], Khalid Alsaleh [c]

[a] Nanomedicine Research Unit, Department of Pharmaceutics, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia
[b] King Abdullah Institute for Nanotechnology, King Saud University, Riyadh, Saudi Arabia
[c] Oncology Center, King Saud Medical City, King Saud University, Riyadh, Saudi Arabia

ARTICLE INFO

Article history:
Received 15 December 2016
Accepted 30 May 2017
Available online xxxx

Keywords:
Lung cancer
Chlamydia pneumoniae
Etiology
Systems biology
Protein targeting

ABSTRACT

Collecting evidence suggests that the intercellular infection of *Chlamydia pneumoniae* in lungs contributes to the etiology of lung cancer. Many proteins of *Chlamydia pneumoniae* outmanoeuvre the various system of the host. The infection may regulate various factors, which can influence the growth of lung cancer in affected persons. In this *in-silico* study, we predict potential targeting of *Chlamydia pneumoniae* proteins in mitochondrial and cytoplasmic comportments of host cell and their possible involvement in growth and development of lung cancer. Various cellular activities are controlled in mitochondria and cytoplasm, where the localization of *Chlamydia pneumoniae* proteins may alter the normal functioning of host cells. The rationale of this study is to find out and explain the connection between *Chlamydia pneumoniae* infection and lung cancer. A sum of 183 and 513 proteins were predicted to target in mitochondria and cytoplasm of host cell out of total 1112 proteins of *Chlamydia pneumoniae*. In particular, many targeted proteins may interfere with normal growth behaviour of host cells, thereby altering the decision of program cell death. Present article provides a potential connection of *Chlamydia pneumoniae* protein targeting and proposed that various targeted proteins may play crucial role in lung cancer etiology through diverse mechanisms.

1. Introduction

Lung cancer is one of the main reasons for cancer-related mortalities world-wide (Siegel et al., 2016). The association of *Chlamydia pneumonia (C. pneumonia)* infections with lung cancer etiology has been previously suggested (Byrne and Ojcius, 2004; Laurila et al., 1997). However, the process of carcinogenesis is not well understood yet. *C. pneumoniae* is a widespread intracellular respiratory bacterium that needs host cell for their existence and multiplication within host cells. It has evolved various effector proteins and toxins that interfere with cell death signalling pathways in host cells and the machinery of cell death executioner (Bohme and Rudel, 2009). Nucleus, mitochondria, and cytoplasm are critical components for normal host-cell functions (Bhavsar et al., 2007; Hess et al., 2003). Mitochondria, which are generally considered to originate from endosymbiotic event, plays an important role in numerous biochemical cascades that direct to programmed cell death (Gray et al., 1999). Therefore, investigating the mechanisms by which intracellular pathogens such as *C. pneumoniae* may alter apoptotic pathways is of special interest.

We have recently studied the nuclear targeting of *C. pneumoniae* proteins through computational approach during intracellular infection where they may play crucial role in progress and development of lung cancer (Khan et al., 2016a). Nevertheless, the fact that many *C. pneumoniae* proteins are targeted into mitochondria, cytoplasm, and other host intracellular components may alter many biochemical pathways in the host. Here we systematically predict several *C. pneumoniae* proteins that can localize into host

* Corresponding authors at: Nanomedicine Research Unit, Department of Pharmaceutics, College of Pharmacy, PO Box 2457, King Saud University, Riyadh 11451, Saudi Arabia.
E-mail address: aalshamsan@ksu.edu.sa (A. Alshamsan).
[1] These authors contributed equally to this work.
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jsps.2017.05.007
1319-0164/

Please cite this article in press as: Alshamsan, A., et al. Prediction of *Chlamydia pneumoniae* protein localization in host mitochondria and cytoplasm and possible involvements in lung cancer etiology: a computational approachChlamydia pneumoniae protein localization in host mitochondria and cytoplasm –>. Saudi Pharmaceutical Journal (2017), http://dx.doi.org/10.1016/j.jsps.2017.05.007

mitochondria and cytoplasm where they can influence the development and progress of lung cancer.

2. Materials and methods

2.1. Protein database for the selection of C. pneumoniae strain

The Uniprot database was screened out for the selection of *C. pneumoniae* strain. Five strains *C. pneumoniae* are publicly accessible in Uniprot database with different number of proteins (Kalman et al., 1999; Myers et al., 2009; Read et al., 2000; Shirai et al., 2000). The *C. pneumoniae* pathogen is reported as obligate gram negative bacteria which may act as a potential factor of lung cancer (Yang et al., 2003; Zhan et al., 2011).

2.2. cNLS mapper for the analysis of NLS

The bioinformatics software cNLS mapper was utilized for the analysis of nuclear localization signal (NLS) in whole proteins of *C. pneumoniae* TW-183 strain (Kosugi et al., 2009a). The whole sequence of all proteins was utilized for the analysis of monopartite and bipartite NLS sequence cutoff value.

2.3. BaCelLo for the analysis of sub cellular protein targeting in different cell compartments

Protein targeting of *C. pneumoniae* TW-183 stain in different eukaryotic cell organelles were analysed by bioinformatics software Balanced Sub Cellular Localization predictor (BaCelLo). The BaCelLo software based on three well known datasets for eukaryotic kingdoms which includes plants, fungi, and animals (Pierleoni et al., 2006). The software analyses the five different classes of sub cellular localization such as mitochondrial, cytoplasmic, nuclear, chloroplast and secretory. The current study was performed with animal datasheet related tool.

3. Results

3.1. Protein database for the selection of C. pneumoniae strain

Whole proteins of *C. pneumoniae* (TW-183 strain) were analysed in present study. We selected *C. pneumoniae* TW-183 strain for our prediction study because of its contains the largest proteome, total of 1112 proteins, as compared to the other four *C. pneumoniae* strains J138, CWL029, AR39 and LPCoLN (Kalman et al., 1999; Myers et al., 2009; Read et al., 2000; Shirai et al., 2000). Scheme showed the summary of work plan for data analysis and possible outcome. Detailed description was presented in the experimental, result and discussion section (Fig. 1).

3.2. Prediction of mitochondrial targeting of C. pneumoniae proteins

The prediction results of BaCelLo have shown 183 proteins target to host cell mitochondria out of total 1112 proteins. It was found that raising bipartite NLS cutoff value of proteins was associated with high probability to targeting in mitochondria. On the contrary, an opposite pattern was noticed with monopartite NLS value of predicted proteins (Table 1). However, no correct connection was predicted with molecular weight and protein targeting to mitochondrial except in few cases where the proteins with minimum molecular weight (0–20 kDa) were mainly localized in host-cell mitochondria (Table 2). The results of pI values were not indicating any constant pattern for targeting of *C. pneumonia* proteins in host mitochondria (Table 3). Multi-modality of pI distribution is a common feature in different whole proteomes. Some researchers studied pI value corelate to the protein localization in different subcellular compartments (Wu et al., 2006). The outlines of proteins localization in host mitochondria illustrated in Fig. 2. Furthermore, as shown in (Table S1), we have provided details about 183 proteins predicted to target mitochondria during our study.

3.3. Prediction of cytoplasmic targeting of C. pneumoniae proteins

The prediction results of BaCelLo have shown 513 proteins targeted to host cell cytoplasm out of total 1112 proteins of whole *C. pneumoniae* proteome. Most of the proteins containing NLS cutoff value of 3–5 were targeted to the cytoplasm with respect to monopartite NLS and bipartite NLS cutoff values (Table 1). Also, most proteins with higher molecular weights were associated host-cell cytoplasmic localization with few exceptions (Table 2). On the other hand, it was observed that the increasing the pI value of the predicted proteins constantly decreased the localization in cytoplasm of host cell (Table 3). The localization patterns of *C. pneumoniae* proteins in host cytoplasm are illustrated in Fig. 3 with different parameters. Moreover, additional data provides the description regarding various *C. pneumoniae* proteins targeting to cytoplasm of host cell are listed in (Table S2).

4. Discussion

Various bacterial proteins modulate host cell survival leading to dynamical suppression of host cell death as a mean for bacterial persistence and multiplication (Fan et al., 1998; Fischer et al., 2001; Rajalingam et al., 2001; Ballestar and Esteller, 2005. *In silico* analysis of protein subcellular localization can significantly help to explain bacterial-protein potential functions. Although many advanced experimental high-throughput approaches have been developed to determine proteins localization, they are time consuming and cost ineffective (Huh et al., 2003; Khan et al., 2016b,2016c; Laurila and Vihinen, 2009; Mote and Reines, 1998). In the advancement of bioinformatics, fast and highly-accurate genome-scale computational predictors of subcellular protein targeting/localization offer a good complement for experimental practice (Garg et al., 2005; Reinhardt and Hubbard, 1998). cNLS Mapper predicts classical NLS functionality of proteins by analysing summation of the functional role of each amino acid in the query protein as per the activity-based profiles, which are achieved from the systematic residue-replacement analyses in *Saccharomyces cerevisiae*. Although analysis of NLS in a particular sequence of protein is essential for analysing its nuclear localization with the highest cNLS score, the cNLS Mapper is a predictor to analyse NLS activity rather than NLS sequence as it consequently predicts cytoplasmic localization of proteins (Hahn et al., 2008; Kosugi et al., 2009a, 2009b).

In contrast, BaCelLo is another predictor that utilizes a decision tree which was developed on diverse SVMs for the prediction of nuclear, mitochondrial, chloroplast, and cytoplasmic targeting of particular proteins (Pierleoni et al., 2006). Also, BaCelLo predicts complete protein sequence along with its N and C termini and analyse the results on the basis of information gained by amino acid residue sequence and evolutionary alignment. To impede different mitochondrial functions, bacterial proteins should attach to- and enter the mitochondria during the course of infection (Cossart and Sansonetti, 2004; Kozjak-Pavlovic et al., 2008). In addition to several anomalous mitochondrial functions, cancer cells have impaired oxidative phosphorylation as a result of the high-rate modified glycolysis with elevated fermentation of lactic acid, which permits cancer cells to retain biosynthetic fluxes at the time of fast proliferation (DeBerardinis et al., 2008; Rudel et al., 2010;

Please cite this article in press as: Alshamsan, A., et al. Prediction of *Chlamydia pneumoniae* protein localization in host mitochondria and cytoplasm and possible involvements in lung cancer etiology: a computational approachChlamydia pneumoniae protein localization in host mitochondria and cytoplasm –>. Saudi Pharmaceutical Journal (2017), http://dx.doi.org/10.1016/j.jsps.2017.05.007